

AD A 095665

RADC-TR-80-342
Final Technical Report
November 1980



UNIVERSAL TRANSPARENT SIMULATOR (UNITRANS)

Pattern Analysis and Recognition Corporation



Eugene H. Lee N. Shohara T. Robacher Mark Nelson, Ph.D.



APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED



ROME AIR DEVELOPMENT CENTER
Air Force Systems Command
Griffiss Air Force Base, New York 13441

81 2 27 007

This report has been reviewed by the RADC Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

MANUFACTURE TO

Some of the machine-produced data are of poor quality reproduction. However, this is insignificant because the relevant information is sufficiently clear to the reader.

RADC-TR-80-342 has been reviewed and is approved for publication.

APPROVED:

JAMES J. MAIER Project Engineer

Yourk, Lunte

James J. Maier

APPROVED:

OWEN R. LAWTER, Colonel, USAF Chief, Intelligence & Reconnaissance Division

FOR THE COMMANDER: John S. Huse

JOHN P. HUSS Acting Chief, Plans Office

If your address has changed or if you wish to be removed from the RADC mailing list, or if the addressee is no longer employed by your organization, please notify RADC (IRAA) Griffiss AFB NY 13441. This will assist us in maintaining a current mailing list.

Do not return this copy. Retain or destroy.

(14) PAR-80-43

390101

SECURITY CLASSIFICATION OF THIS PAGE(When Date Entered)

Item 20 (Cont'd)

cont.

>testing device which could impact a variety of electronic signals to personnel identification and authentication equipments, and obtain a response which could be used as a substitute for the testing by entrants. The work of the contract demonstrated that the techniques available need much further development work.

Acces	ssion For	
NTIS	GRA&I	×
DTIC	TAB	
Unannounced 🗍		
Justification		
Ву		
Distribution/		
Availability Codes		
	Avail and	l/or
Dist	Special	L
	1 1	
	1 1	
H	1	
	1	

UNCLASSIFIED

TABLE OF CONTENTS

Section		Page
1.	Introduction	1-1
1.1.	Objectives	1-1
1.2.	Executive Summary	1-3
1.3.	Organization	1-4
2.	Speaker Simulation	2-1
2.1.	Introduction	2-1
2.2.	General Concept	2-1
2.3.	Linear Predictive Coding (LPC) Its Rudimentary Algorithm	2-3
2.4.	PAR Speech Processing (PSP) Capability[8]	2-4
2.5.	Simulated-Speaker Phoneme Library Software Design	2-4
2.6.	Parameter Construction	2-8
2.7.	Discussion of Algorithms Used in Speaker Simulation.	2-9
2.8.	Experimental Results	2-18
3.	Fingerprint Simulation	3-1
3.1.	Introduction	3-1
3.2.	Fingerprint Pattern Classification	3-1
3.3.	Generation of Fingerprint Image	3-4
3.4.	Typical Examples	3-7
3.5.	Further Development	3-12
4.	Signature Simulation	4-1
4.1.	Introduction	4-1
ir o	Models for Signatures	4-2

TABLE OF CONTENTS CONT'D

Section		Page
5.	Comparison of UNITRANS with VADABS System	5-1
5.1.	Introduction	5-1
5.2.	VADABS System and its Limitation	5-1
5.3.	UNITRANS Approach	5-4
6.	Recommendations	6-1
	References	R-1

LIST OF FIGURES

Figure	Pag	<u>e</u>
2-1	Speaker Simulation and its use for Device Testing . 2-2	!
2-2	Block Diagram of Simplified Model for Speech Production	j
2-3a	PAR Speech Processing (PSP) System 2-6	i
2-3b	Block-Diagram for Speaker Simulation 2-7	
2-4	Transition Handling During Concatenation 2-1	.0
2-5	Distribution of Area Segment i over Real Sample Data and the Estimated Population Distribution Ppop(A ₁)	.1
2-6	Synthetic Speech = 'North' for Simulated Speaker 'SPAA'	:0
2-7	Synthetic Speech = 'North' for Simulated Speaker 'SPBB'	1
2-8	Synthetic Speech = 'North' for Simulated Speaker 'SPCC'	2
3-1	Fingerprint Patterns of the Henry System 3-3	
3-2	Impose Minutiae on Fingerprint Pattern 3-6	
3-3	Physical Fingerprint Generation Scheme 3-8	
3-4	Regions Assignment for Fingerprint Simulation 3-1	2
3-5	Loop Type Fingerprint Simulation	3
3-6	Arch Type Fingerprint Simulation	5
3-7	Whorl Type Fingerprint Simulation	7
3-8	Copper Clad Board With Fingerprint Images Taken in Different Exposures	9
4-1	Five Repetitions of a Signature by the same Individual	
4-2	Signature of Ludwig II of Bavaria, which Contains a Swan, Symbol of His House 4-5	

LIST OF FIGURES CONT'D

Figures		Page
4-3a	Five Signature Repetitions - Eyes Open	4-6
4-3b	Five Signature Repetitions - Eyes Closed	4-6
4-4a	Five Signature Repetitions - Eyes Open	4-7
4-4Ъ	Five Signature Repetitions - Eyes Closed	4-7
4-5a	Five Signature Repetitions - Eyes Open	4-8
4-5b	Five Signature Repetitions - Eyes Closed	4-8
4-6	Signatures may be Encoded as a Sequence of Velocity Vectors	4-10
4-7	Letter 'M' x and y Curves	4-11
4-8	The Trajectory Vectors of the Letter M plotted in Polar Form	
4-9	Letters Written According to Palmer System	4-16
5-1	VADABS System Concept	5-2

EVALUATION

Contract No.: F30602-79-C-0226, "Universal Transparent Simulator (UNITRANS)"

The Universal Transparent Simulator (UNITRANS) effort is part of the Center effort to develop improved Entry Control Systems under TPO 5, and was investigated to develop a substitute for the use of many entrants to test entry control systems.

The work performed under this effort demonstrated that the techniques which are available need much further development work, making the whole concept of a universal tester more complex than originally envisioned. This was especially true of the fingerprint and the handwriting systems. There was some success in the development of synthetic speech.

There is no further work in this area planned at this time.

JAMES J. MAIER

Project Engineer

1. INTRODUCTION

1.1 OBJECTIVES

Over the past two decades a considerable amount of research has been expended in the field of pattern recognition. Examples of automatic pattern recognition systems exist in abundance. Some focus on commercial applications, while others on military and high-security operations. Access control is the high-security process that limits secure area entry to authorized individuals. In order for these people to gain access to a secured location, some type of personal information is used for identification. Basically, three methods of confirmation of identity are used for individual verification: (1) something "known" by the individual, such as password and combination of a lock; (2) something "possessed" by the individual, such as badges, passes and keys; and (3) something "about" the individual, such as individual's appearance, voice and fingerprints. Because of the vulnerability of the first two categories to such threats as theft and duplication, much presently focused on personal identification physiological attributes. (Table !-1 compares the degree of security level for personal identification.) [1]

Rome Air Development Center (RADC) is currently undertaking tests on those access control devices that identify human beings based on personal attributes. Existing devices measure voice, fingerprint and signature. Potential devices may measure palmprints and c-trace (a type of electrocardiogram) [2]. In order to test these devices with reliable results, a large data base is necessary. However, data collecting is usually a time-consuming and laborious procedure. Money may be wasted on unreliable data due to noise interference, equipment failure, or human error. In addition, the lost data is not easily recovered due to environmental considerations,

Security Level	Requirement For Security Breach	Example Basis For Identification
Minimum	No Forgery Needed	o Code-Only System o Household Quality Key System
	Straight Forward	o Card System
	Amount of Forgery	o Sophisticated Key System
		o Card-Code System
	Readily Duplicatable	o Personal Appearance
	Personal Attributes	System
		o Hand Geometry System
	Difficult to Duplicate	o Voice Print System
V	Personal Attributes	o Fingerprint System
Maximum	Presently Non-Duplicatable Personal Attributes	o Genetic Code System

TABLE 1-1:
The Spectrum of Security Levels For
Personal Identification [1]

considerations, financial support, and so forth.

This final report describes a method, developed under contract UNITRANS (Universal Transparent Simulator), to simulate personal attributes using a computer so as to generate a large data base from a relatively small sample of actual data. The procedure involves generating a reasonable set of features and exploring a way to convert these features to a simulated phenomenon. Two kinds of personal attributes have been simulated and are described in this report -- voice print and fingerprint. Signatures were not generated, but the technique of simulation will be discussed briefly.

1.2 EXECUTIVE SUMMARY

The attributes considered in this research were voice print, fingerprint, and signature. Although the general philosophy and approach are the same in case. the requirements on the aystem necessitate implementations. Speaker simulation is generated by distribution sampling on vocal tract area functions using the Monte Carlo technique. Area functions are used due to their physiological significance and ability to represent the statistical information across the population. A phoneme library is built for an individual simulated speaker so that an unlimited vocabulary can be generated. During the construction of words from phonemes, rate of change of area components are assigned during the transition period. The digitized speech signal is then decoded back to physical voice by using the LPC synthesizing filter and then used as an input to the devices for testing. The result comes out rather satisfactorily. For illustrative purposes, the word 'north' was simulated. First, a simulated phoneme library was built for a simulated speaker. A phoneme string was then constructed by selecting the desired phonemes from the library. For the word "north", phonemes /NN/, /OY1/, /OY2/, /TH/, and /SI/ were used to build up the string. By specifying the pitch %, gain %, and phoneme duration, a continuous speech signal was reconstructed. The simulated synthetic speech 'north' came out very intelligible, and the simulated speakers generated from different random

number seeds were also distinct from each other.

In the area of fingerprint, three types of patterns have been simulated so far. They are loop, arch and whorl. A basic pattern is built for each type initially, then by applying the topological transformation to each pattern, minutiae such as ridge endings and bifurcations can be added to any part of the print the operator desires. Once a satisfactory digital image is formed, photo-etching followed by molding produces a 'live' fingerprint.

Signature has not been studied in detail in this research. One approach using (X,Y) directions as a function of time, pressure, slope, curvature, etc. is recommended for features representing signature. Then by applying the Monte Carlo technique on the distribution of the parameters, individual signatures could be simulated.

1.3 ORGANIZATION

The presentation of this research is organized as follows:

Section 2 is concerned with a detailed development of a speaker simulation algorithm. The distribution sampling via Monte Carlo technique and the method of phoneme reconstruction during transition period is introduced. The first generation PAR Speech Processing (PSP) system is illustrated, followed by the description of how this system is used to create the simulated-speaker phoneme library.

In Section 3, fingerprint simulation technique is introduced. The imitation of fingerprint image via topological transformations is first presented, then followed by the description of photoetching method for `live' fingerprint generation.

Section 4 gives a brief suggestion on signature simulation. Some basic approaches are described and illustrated.

Section 5 gives a discussion of the proposed VADABS system, the goal of which is identical to UNITRANS. The comparison of this system with our UNITRANS system is presented. The limitations of VADABS and our different approach on UNITRANS are also described.

Finally, Section 6 gives the recommendations of this work.

2. SPEAKER SIMULATION

2.1 INTRODUCTION

Speech synthesis and voice recognition are receiving increasing attention these days as a growing number of consumer and commercial products rush to the market with low cost and claims of high performance. According to SRI's report on the speech communications equipment market, speech recognition equipment will account for \$995 million to \$1.455 billion with an average annual growth rate of over 50% by 1988, and the current aerospace and defense market should add one quarter to one-half again to the total for all other markets [3-4]. Since speech recognition systems are increasing and most of them claim to achieve more than 95% accuracy, there is a growing concern about how to evaluate, especially how to collect a large speech data set for such evaluation.

2.2 General Concept

A general block diagram on the PAR concept of speaker simulation and approach to the laboratory testing of an access control device is shown in Figure 2-1. Basically, the speaker simulation scheme is based on the distribution of vocal tract cross-sectional areas encoded from a small group of people. A large data set of simulated area functions can then be generated by sampling from the distribution. For the purpose of generating physical human voice, area functions are decoded back to speech signals using the linear predictive coding (LPC) technique (a full description on LPC will be introduced in the next section.) To simulate an individual successfully, the distribution of day-to-day variations must also be taken into consideration. hus, data collection proceeds in two steps: (1) on a person-to-person basis, and (2) on a day-to-day basis for each person.

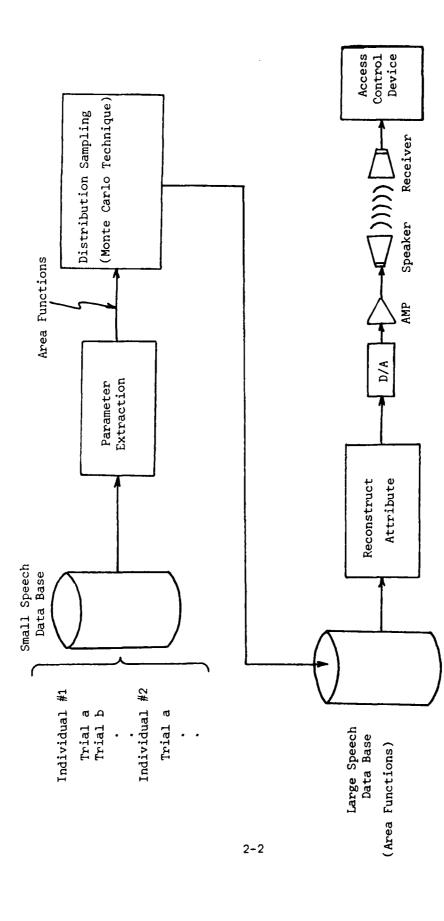


Figure 2-1 Speaker Simulation and its use for Device Testing

2.3 LINEAR PREDICTIVE CODING (LPC) -- ITS RUDIMENTARY ALGORITHM

LPC is a speech synthesis technique that allows a reproduction of original vocal sounds by means of relating one voice signal sample linearly to its neighboring sample. Suppose X(n) is the signal value at time n, then the predicted value X(n) can be represented by

$$\begin{array}{ll}
\ddot{X}(n) &= \sum_{i=1}^{M} a_i X(n-i) \\
\dot{i} &= 1
\end{array} (2.1)$$

where a 's are LPC coefficients to be found.

In this equation, since X(n) is the approximation to X(n), it is possible to estimate the associated error E(n), alternatively referred to as the residual:

$$E(n) = X(n) - \tilde{X}(n)$$
 (2.2)

The linear prediction parameters can then be obtained when the value of E(n) becomes minimum. Because E(n) also serves to provide information on voice sources, it is sometimes referred to as the driving function.

The X(n) can be reproduced by feeding into a digital filter the transmission characteristics. This synthesized filter is implemented based on a series of cascaded acoustic tubes. Such a model and its relation to linear prediction coefficients is described in detail by Markel and Gray [5]. The vocal tract transfer function H(z) is written as

$$H(z) = \frac{1}{M}$$

$$1 - \sum_{i=1}^{N} a_i z^{-1}$$
(2.3)

and the general synthesizing process is illustrated in Figure 2-2 below.

2.4 PAR SPEECH PROCESSING (PSP) CAPABILITY[8]

A PAR Speech Processing (PSP) System based on the LPC technique has been developed (Figure 2-3a). This system permits a user/analyst to digitize or playback, encode or decode speech signals. Each function in the system is carried out by a separate task. The tasks communicate by passing data through data files and execution parameters through a parameter file. The interface to the user is handled by an indirect command file processor routine which prompts the user for a command, executes the desire task, then returns for another command.

In the speech synthesis process, basic speech sounds, known as phonemes, are extracted from a sample of the subject's speech, then used to construct an arbitrary utterance. The phonemes are stored in a library as parameter sets which define a state of the vocal tract model: these parameters are the area coefficients, voicing, pitch, gain, LPC and reflection coefficients. On construction, parameter sets are allowed to change at pre-defined rates from one state to the next, but in each state, all parameters are fixed. Current capabilities include: encoding and decoding of speech signals, display of waveforms and area functions, creation of and addition to phoneme libraries, and specification and construction of a synthetic utterance.

2.5 SIMULATED-SPEAKER PHONEME LIBRARY SOFTWARE DESIGN

A block diagram for creating a simulated-speaker phoneme library is shown in Figure 2-3b. Input speech signals are digitized by using 12.8 KHz sampling rate. Because of the non-stationary properties of speech, they are encoded on a frame-by-frame basis with 20 ms per frame. After segmentation, encoded parameters are stored in a file. Among those encoded parameters, vocal tract cross-sectional areas are used as attributes for speaker simulation since they can provide a reasonable physiological meaning for speech production. In

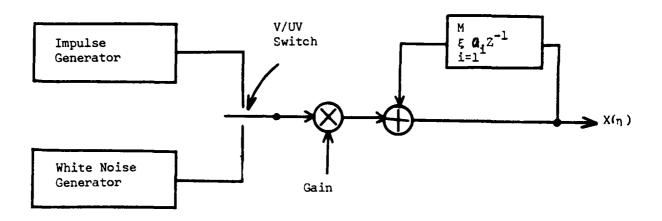


Figure 2-2: Block Diagram of Simplified Model for Speech Production

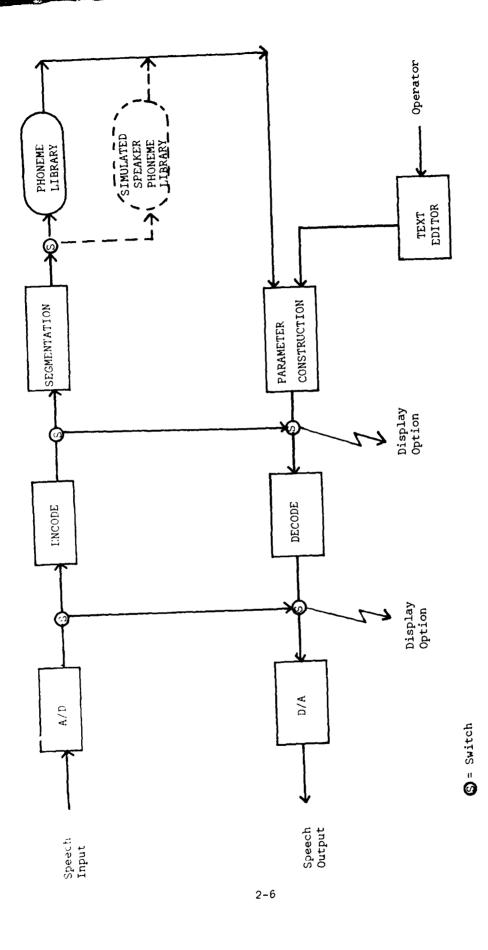


Figure 2-3a PAR Speech Processing (PSP) System

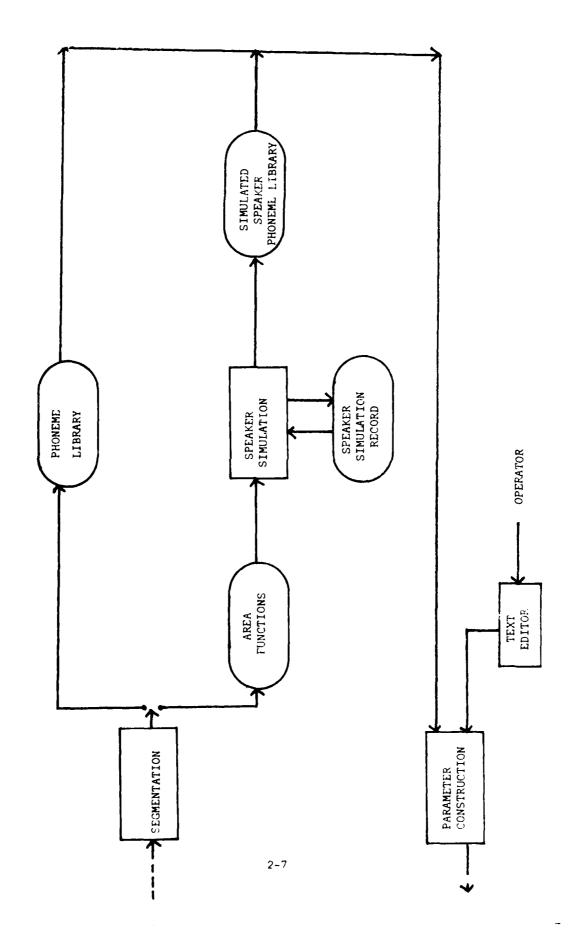


Figure 2-3b Block-Diagram for Speaker Simulation

addition, the population and individual distributions of area functions are more easily interpreted and understood than are LPC or reflection coefficients.

Each simulated speaker is then represented by a vector of 15 area functions coefficients for each phoneme sampled from the population distribution. This information is stored in a simulated-speaker phoneme library under the simulated speaker name. A simulated-speaker file is also created that holds several records containing information such as speaker names, phoneme names, means and variances. In the process of simulation, an operator will enter a command string consisting of simulated speaker name and phoneme name. The computer will check the simulated-speaker record to see if the requested speaker and phoneme is already in file. If not, the required entry will be created. The simulated speech signal can then be constructed by activating the "parameter construction" program which will be discussed in Section 2.6, and the simulated-speaker or phoneme can be listened to, displayed, or used as the operator desires.

2.6 PARAMETER CONSTRUCTION

In the simulation of spoken phrases concatenated from simulated phonemes, rates of change of area functions are assigned during the transition period. The rate is assigned so that for any particular area component changes are directly proportional to the maximum range over which the area must vary. Suppose R_{i}^{α} is the rate of change of area component A_{i} of speaker α between phoneme 1 and phoneme 2, then the rate at which area component A_{i} changes is given by:

$$R_{\mathbf{i}}^{\alpha} = C(A_{\mathbf{i}}^{\alpha 1} - A_{\mathbf{i}}^{\alpha 2}) \tag{2.4}$$

where C is constant.

With the rates assigned we would then envision a time sequence t_{1i} , t_{2i} , t_{3i} for the concatenation of phoneme 1 and phoneme 2.

Suppose t_{1i} and t_{3i} represent the duration of two phonemes to be concatenated (see Figure 2-4), and let

$$t_{2i} = \min(t_{1i}, t_{3i})$$
 (2.5)

be the transition interval. Equation (2.4) can be rewritten as:

$$R_{i}^{\alpha} = \frac{1}{t_{2i}} \left(A_{i}^{\alpha 1} - A_{i}^{\alpha 2} \right) \tag{2.6}$$

Similarly, the rate of change of pitch period and gain during the transition interval can be estimated by applying Equation (2.6) with area components replaced by pitch periods and gains.

2.7 DISCUSSION OF ALGORITHMS USED IN SPEAKER SIMULATION

2.7.1 Basic Approach

In order to simulate a large population with only a few real sample data at hand, we assume that the overall distribution is Gaussian. Such an assumption is reasonable provided that the data collected are within a group. That is to say, three different kinds of data bases will be collected so as to estimate three different sets of distributions -- namely, male, female, and child. The relationship between the subsets of the data base and overall population for area A_i is suggested in Figure 2-5. The values of the individual sample population are used to estimate the overall distribution $P_{pop}(A_i)$. Of course, the ability to resolve details in $P_{pop}(A_i)$ depend on

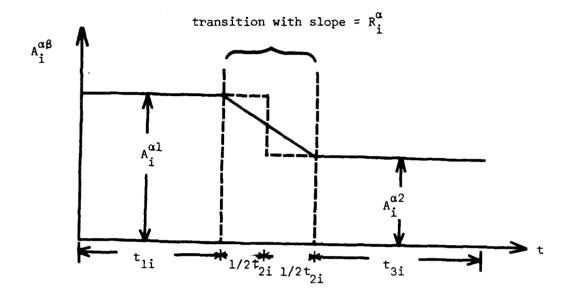


Figure 2-4 Transition Handling During Concatenation

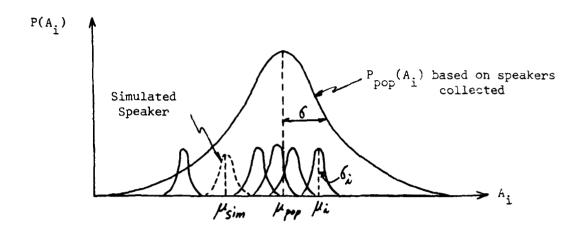


FIGURE 2-5. Distribution of Area Segment i over Real Sample Data and the Estimated Population Distribution $P_{pop}(A_i)$.

the number of samples available.

From Figure 2-5, assume i is the i-th area segment for speaker ω with mean μ_i and variance σ_i . Given a reasonable data sample for speaker ω it is always possible to obtain the distribution p_i (A_i). With a group of speakers available, the distribution for this group can be estimated.

if the data base collected happens to be really sparse, for instance only one speaker is collected, $P(A_{\underline{i}})$ estimated may not be sufficient to represent the overall population. One solution is to artificially produce a population distribution by making the variance of $P(A_{\underline{i}})$ several times larger, i.e.,

$$\sigma_{\text{pop}} = (\text{factor}) (\sigma)$$
 (2.7)

provided that the area distribution is Gaussian. Actually, in the absence of a large sample data base, the assumption of a simple analytic form (Gaussian) for the distribution is reasonable. If several speakers are available, such a magnification procedure will not be necessary.

The mean $\mu_{\mbox{sim}}$ for a simulated speaker now can be generated by

$$u_{sim} = \mu_{pop} + x\sigma_{pop}$$
 (2.8)

where x is a Gaussian random number.

Likewise, the variability of vocal tract area functions A within the simulated speaker can be estimated by

$$A_{i} = \mu_{sim} + y_{\sigma_{avg}}$$
 (2.9)

where y is another Gaussian random number different from x, and $\frac{\sigma}{\text{avg}}$ is the average variance such that $\frac{\sigma}{\text{avg}} = E \left\{ \frac{\sigma}{i} \right\}$. That is, $\frac{\sigma}{\text{avg}}$ is the typical variability within a single speaker.

2.7.2 Further Development for Correlated Area Functions

In the foregoing discussion, we have made the assumption that the area functions are uncorrelated with each other. In fact, that is not true. The development, however, can be easily extended to the case when correlated data are present.

Suppose i and j denote the i-th and j-th segment of vocal area, a and a j. Then the covariance σ_{ij} and mean values μ_i for k utterances of phoneme β by a particular speaker can be computed as

$$\mu_{i} = \overline{a}_{i} = \frac{1}{k} \sum_{k=1}^{k} a_{i}(k)$$

$$\sigma_{ij}^{2} = E\{(a_{i} - \overline{a}_{i}) (a_{j} - \overline{a}_{j})\}$$

where E represents the expected value operator.

A more general notation can be developed. We have 15 vocal tract segments (i.e., i=1,2,...15). Let A_{ω}^{T} (m) = $[a_{1}(m)a_{1}(m)...a_{15}(m)]$ be the transpose of an area vector for the ω -th speaker on the m-th utterance of phoneme β . Then the covariance matrix of A_{ω} is defined as

$$\Sigma_{\omega} = E\{(A_{\omega} - \overline{A}_{\omega})(A_{\omega} - \overline{A}_{\omega})^{T}\}$$

$$= E\{A_{\omega}A_{\omega}^{T}\} - \overline{A}_{\omega}\overline{A}_{\omega}^{T}$$
(2.10)

If we have ω speakers, then the average covariance matrix Σ is given by

$$\Sigma_{\text{avg}} = P_1 \Sigma_1 + P_2 \Sigma_2 + \dots + P_{\omega} \Sigma_{\omega}$$
 (2.11a)

where " ω " is the number of speakers and Σ_{ω} denotes the covariance matrix of the ω -th speaker with a priori probability P_{ω} , and the overall covariance matrix is

$$\Sigma = E\{AA^{T}\} - \overline{A} \overline{A}^{T}$$
 (2.11b)

where

$$A = \{A_1, A_2, \dots, A_{\omega}\}$$

Or, Equation (2.11b) can be expressed as

$$\Sigma = \begin{bmatrix} \sigma_{1,1}^{2} & \sigma_{1,2}^{2} & \cdots & \sigma_{1,15}^{2} \\ \sigma_{2,1}^{2} & \sigma_{2,2}^{2} & \cdots & \sigma_{2,15}^{2} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{15,1}^{2} & \sigma_{15,2}^{2} & \cdots & \sigma_{15,15}^{2} \end{bmatrix}$$
(2.12)

2.7.2.1 Rotational Transformation

We want to generate a distribution which has covariance matrix Σ . One way to accomplish this is to transform Σ to its diagonal form, generate 15 independent random numbers, and transform back. Such a transformation is equivalent to a rotation of the original pattern space to a new set of coordinate vectors which are aligned with the distribution for A.

Suppose A is a vector of zero mean, and T is such a linear transform function that maps A to A', then the transformation results in a covariance function that is diagonal with entries equal to the eigenvalues of T. That is, if

then
$$\Sigma' = \mathbf{T}\mathbf{A}$$

$$\Sigma' = \mathbf{E} \{ (\mathbf{T}\mathbf{A})(\mathbf{T}\mathbf{A})^{\mathrm{T}} \}$$

$$= \mathbf{E} \{ \mathbf{T}\mathbf{A}\mathbf{A}^{\mathrm{T}} \mathbf{T}^{\mathrm{T}} \}$$

$$= \mathbf{T}\mathbf{E}(\mathbf{A}\mathbf{A}^{\mathrm{T}})\mathbf{T}^{\mathrm{T}}$$

$$= \mathbf{T}\mathbf{\Sigma}\mathbf{T}^{\mathrm{T}}$$

$$= \mathrm{diag} (\lambda_{1}, \lambda_{2}, \dots, \lambda_{15})$$

That is, the coefficients in the rotated space become statistically uncorrelated. Thus, if $A^T = [a_1, a_2, \ldots a_{15}]$ is normally distributed with covariance matrix Σ , then the rotated feature vector A' = TA has the property that each of its components a_i is normally distributed with variance λ_i . Thus, A may be sampled by sampling each a_i independently and making the transformation $A = T^{-1}A = T^TA$ back to pattern space. However, this transform function, which is usually called the Karhunen-Loéve Transform (KLT), cannot be implemented efficiently due to lack of a general algorithm that enables its fast computation. Many other rotational transforms have therefore been used

by researchers, providing almost uncorrelated data in feature space. These transforms are readily implementable in a digital environment. The most commonly used transforms are fast Fourier transform, discrete Cosine transform, Haar transform and Walsh-Hadamard transform.

2.7.2.2 'P' Transform Function

The above described technique which must use a transformation both forward and backward is computationally inefficient. We shall now describe the technique which we actually employed, the 'P' transform.

For simplicity of illustration, we shall demonstrate the technique for a two-dimensional measurement space. The problem then is to generate a two-dimensional Gaussian distribution with covariance matrix

$$\sum_{\sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}$$
 (2.13)

Let $\mathbf{U}^T = \{\mu_1, \mu_2\}$ be a random vector of zero mean, unit covariance, and uncorrelated components. Let $\mathbf{V}^T = [\mathbf{V}_1, \mathbf{V}_2]$ be the result of multiplying \mathbf{U}^T by the transformation P.

$$V = PU, \qquad (2.14)$$

where P is a lower triangular matrix,

$$P = \begin{bmatrix} P_{11} & 0 \\ P_{21} & P_{22} \end{bmatrix}$$
 (2.15)

We can find the elements of the P matrix by forming the expectation value,

$$E\{VV^{T}\} = E\{(PU)(PU)^{T}\}\$$
 (2.16)

$$= E\{PUU^{T}P^{T}\}$$
 (2.1°)

$$= E\{PP^{T}\} = PP^{T} \tag{2.18}$$

since $\mathbf{W}^{\mathbf{T}}$ is the identity matrix according to our assumption on U. Thus

$$\begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 \\ \sigma_{12}^2 & \sigma_{22}^2 \end{bmatrix} = \begin{bmatrix} P_{11}^2 & P_{11}P_{21} \\ P_{11}P_{21} & P_{21}P_{21} + P_{21}^2 + P_{22}^2 \end{bmatrix}$$
(2.19)

From Equation (2.19), it follows that from the covariance matrix of area functions which we calculate from the data base, the "P" matrix can be determined. With this matrix, simulated speaker vocal tract area functions can be generated by using a Gaussian random number generator. This technique is much more useful and efficient than the others mentioned before, since it can eliminate many transformation steps as far as speaker generation is concerned. This technique can easily be extended to a multidimensional space with the upper triangular terms in "P" set to zero.

Again, in order to estimate the overall population density function if only one speaker is collected for data base, the covariance matrix based on this sample data is multiplied by a factor (of course this step will not be necessary if several speakers are available). That is

$$\Sigma_{\text{pop}} = (\text{factor}) (\Sigma)$$
 (2.20)

and Equation (2.8) and (2.9) become

$$[\mu_{sim}] = [\mu_{pop}] + [X*]$$
 (2.21)

[A] =
$$[\mu_{sim}] + [Y^*]$$
 (2.22)

where

$$[x*] = [P_{pop}][x]$$

 $[Y*] = [P_A][Y]$

and [X], [Y] are two Gaussian random number vectors.

The areas simulated are then transformed into LPC and reflection coefficients using a step-up procedure given in Markel and Gray [5]. All these parameters together with phoneme label, voicing, pitch, and gain are put into the named simulated-speaker phoneme library.

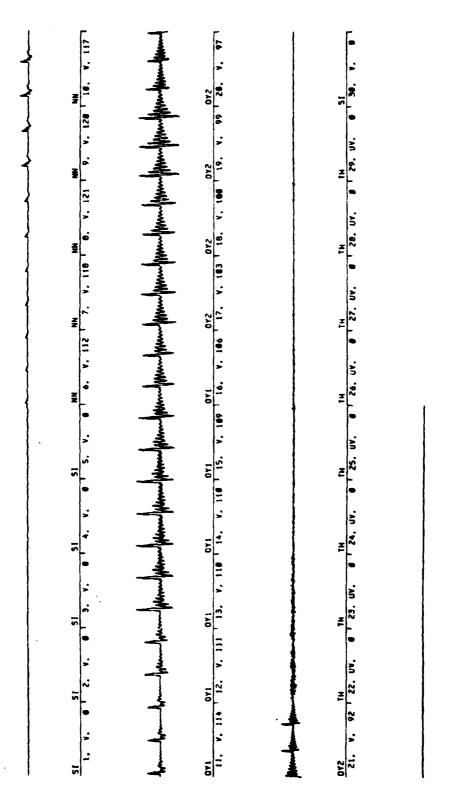
2.8 EXPERIMENTAL RESULTS

In the data collecting process, three male speakers were chosen to read a phrase 'north lawn great camp' five times. The continuous speech waveform was then digitized by using 12.8 KHz sampling rate. The speech signal was then encoded in a frame-by-frame basis with 20 ms of 256 samples per frame. The waveform was then displayed and the desired phonemes were selected. For the purpose of illustration, the word 'north' was chosen for simulation

demonstration. Suppose the word 'north' is represented by /NN/ /OY1/, /OY2/ and /TH/. then the total data base for each phoneme will be 15 frames. By applying the statistical method and random number generator on the area functions described earlier, a simulated phoneme which is represented by 15 area segments was generated. This procedure was repeated for other phonemes, and so a library of simulated phonemes for a particular simulated speaker was created (Table 2-1). The next step is to build a phoneme string for speech reconstruction. By using the text editor, an operator can build a phoneme string by specifying pitch %, gain %, and duration. One example of building such a string for the word 'north' is shown in Table 2-2. Once the phoneme string is set, the simulated synthetic speech signal can be constructed by activating both the 'parameter construction' program and the 'decode' program available in the PSP system, and the simulated speaker or phoneme can be listened to, displayed or manipulated as the operator desires (Figure 2-6).

Other simulated speakers were also generated by using different random number seeds and the results are shown in Table 2-3 and Figure 2-7 for simulated speaker SPBB, TABLE 2-4 and Figure 2-8 for simulated speaker SPCC. When the simulated utterance 'north' is played back through the D/A converter, the speech is surprisingly natural and intelligible, and the difference between the three simulated speakers can also be distinguished.

PILES; WORTH, SMY
TTI -- STOP -- REOF'S DISPLAY COMPLETED



FIGURE, 2-6. Synthetic Speech = 'North' for Simulated Speaker 'SPAA'

FILES, WORTH, SHY
TTE -- STOP -- REOFE DISPLAY COMPLETED

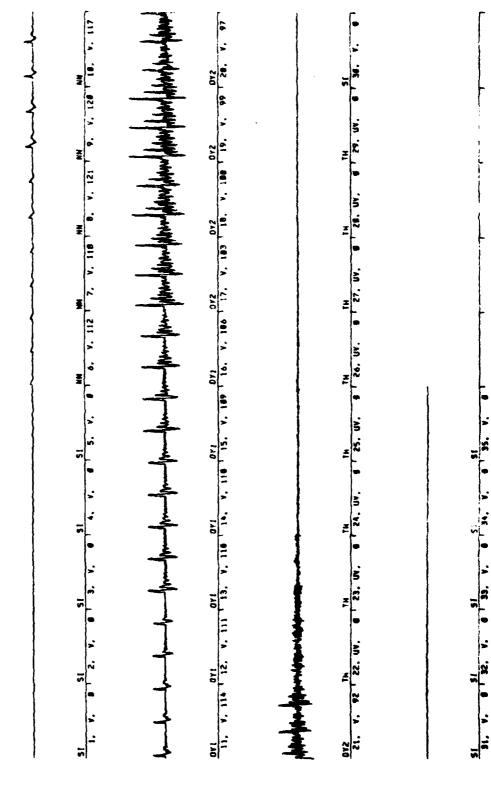


FIGURE 2-7. Synthetic Desch - 'North' for Simulated Speaker 'SPBB'

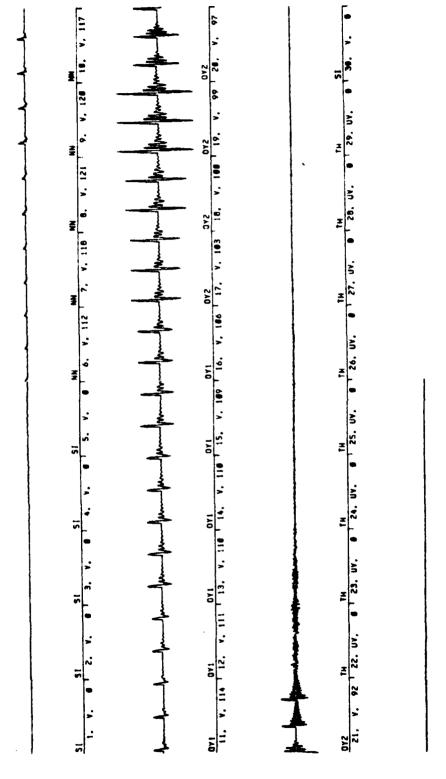


FIGURE 2-8. Synthetic Speech = 'North' for Simulated Speaker 'SPCC'

NEME IS THE CATALOG OF ENTRIES FOR! SPAN.SIM

%%; -	4.0 6.0 70.0 1.1	6.6 6.6 1.63	6.00 9.00 81.00 81.00	€ 6 6 6 6 6 6 7 7 7 8 7 8 8 8 8 8 8 8 8 8	
60.5 7.53 7.63				• = Ø	
	⊕ ⊕ .c.	9.18 1.73	6-6-6 8-6-8	6.00 8.00 8.00 8.00	
	-6.15 9.25 3.58	. 65 8. 65 8. 65		-6.62 6.71	
		44.25 25.15 25.15	-6.21 -6.68 -6.71		
	-6.64 1.45	1.13	4.00 19.00 19.00 19.00	-6.11 -6.14 0.60	
6.4. 1.6.0 1.6.0	6.66 8.9.9 8.18			⊕ ⊕ ₩ ₩ ₩	
 4.1.6	4	1.64	6 6 6 4 6 6 4 6 6	6.00 6.00 8.00 8.00 8.00 8.00 8.00 8.00	
• • • • • • • • • • • • • • • • • • •	 1.43		4.69	-0.19 -0.17 0.20	
	-0.11 0.13 1.87	-0.27 -0.27 -0.71	6. 6. 6. % % %	0.16 0.07 0.23	
6.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00	2.1.8	43 6.33 5.37	£.69	-0.16 -0.38 -0.38 0.10 CATALOG	
	**************************************		9.00 W. W. W. W.	0.18E+02 -0.30 -0.16 -0.43 -0.38 0.04 0.10 LIBRARY CATALO	
121	-1.4 -4.4 -4.4 -4.4	80.08 10.08 13.08 13.08	•k%8	2.33 2.33 2.33 2.33	
1 + 4 + 4 + 4 + 4 + 4 + 4 + 4 + 4 + 4 +	88:	. •••• E&	885	70.00 10.00	
# # # # # # # # # # # # # # # # # # #	9 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	% 4.4.4. 2.6.8.	* 	7,48; 6,46; 5,46;	^

[RECORD NO.]
PH. LABEL, VOICE(T)/UNVOICE(F), PITCH (NO. CF SAMPLES), GAIN
LPC COEFFICIENTS
REFLECTION COEFFICIENTS
AREA FUNCTIONS

NOTE: 'SI' denotes silence (voice with zero pitch)

TABLE 2-1 Phoneme Library for Simulated Speaker = SPAA

Phoneme Label	Pitch %	Gain %	Duration (# of Frame)
SI	0	50	5
NN	100	100	5
OY1	100	100	6
OY2	100	100	5
тн	100	60	8
SI	0	50	6

TABLE 2-2.

Phoneme String for Synthetic Speech = 'North'

The Comp 13	HEADING												
0.05 0.06 0.01 0.01 0.01 0.01 0.01 0.01 0.01	"	•											
1.98 1.11 1.46 2.57 4.66 5.69 4.76 2.46 3.11 1.98 1.11 1.11 1.46 2.59 4.76 2.46 3.11 1.11 1.11 1.11 1.11 1.11 1.11 1.1		4-10	5.15	6.13	96	90 E	9.19	6.67 6.69	-6.14 -6.96	6 .17	-6.67 6.61	6 .17 6 .22	6.13
2.3 110 0.35E+03 -0.50 0.35 0.03 -0.50 0.35 0.03 -0.50 0.35 0.03 -0.50 0.35 0.03 -0.50 0.35 0.03 -0.50 0.03 -0	1.98	1.46	25.52	4.06	5.83	4.76	2.40	3.11	2.60	8.8 8.8	2.15	5. 9 6	 W
-6.56 6.36 6.22 6.35 -6.24 -6.46 6.13 6.65 6.35 6.42 6.93 6.42 6.35 6.35 6.35 6.35 6.35 6.35 6.35 6.35	ري د		:+03					;	į	,	•	1	•
12.87 7.56 9.36 -9.15 4.76 3.56 3.88 2.36 9.97 9.51 3.51 3.56 3.88 2.36 9.97 9.51 3.51 3.56 3.88 2.36 9.97 9.51 3.51 3.56 3.88 2.36 9.97 9.51 9.51 3.51 3.56 3.88 2.36 9.97 9.51 9.51 9.51 9.51 9.51 9.51 9.51 9.51	-1.50		9.35	-0.24	- - - - - - - - - -	6.13	9.65	6.6 6.6	6. de	4. 6.		رة و الأراز	-6.15
93			4.76	3.56		, v.	6.9	6.51	96:	, n	8.35	1.85	6
41 - 6.18 - 6.68 6.18 - 6.69 6.18 - 6.93 6.94 6.38 6.75 - 6.35 6.14 - 6.16 6.52 - 6.11 - 6.12 - 6.13 6.75 6.41 - 6.12 - 6.13 6.75 6.41 - 6.12 - 6.13 6.75 6.41 - 6.12 - 6.13 6.75 6.41 - 6.12 - 6.13 6.75 6.41 - 6.12 6.75 6.41 - 6.12 6.75 6.41 6.12 6.13 6.75 6.13 6.15 6.15 6.15 6.15 6.15 6.15 6.15 6.15	E		E. 64										
6.53 6.75 -6.36 6.14 -6.16 6.52 -6.11 -6.12 -6.13 4.16 6.55 -6.11 -6.12 -6.13 4.16 5.64 1.78 2.29 2.79 4.3 6.29 6.75 6.18 -6.20 6.79 6.18 6.75 6.18 -6.20 6.48 -6.52 6.15 -6.15 6.35 6.35 6.48 -6.15 -6.15 -6.15 6.35 6.35 6.48 -6.15 -6.15 -6.15 6.35 6.35 6.35 6.35 6.35 6.35 6.35 6.3	4		. 4±	-0.66	9.18	-6.93	0.04	9.38	96.	-6.43	9.16	6.13	0 0 0
57.08 17.51 2.54 5.43 4.10 5.64 1.78 6.69 5.73 4.10 5.64 1.78 6.69 5.73 6.10 5.64 1.78 6.69 5.73 6.10 6.10 6.10 6.10 6.10 6.10 6.10 6.10	6 .53		9.14	-0.16	6.52	-0.11	-0.12	-6.13 3.13	9. C	9	7 0 9 4	ัง รัง	
43 6 6.23E+03 1.06 6.75 6.18 -0.89 -0.23 -0.36 -0.48 -0.52 -0.48 1.06 6.75 6.18 -0.09 -0.23 -0.16 -0.08 -0.15 -0.15 6.33 6.47 6.12 6.18 -0.15 6.15 6.19 6.22 6.26 6.36 5.3 6.47 6.15 -0.21 6.14 -0.08 -0.27 -0.05 -0.54 -0.38 -0.15 -0.21 6.14 -0.08 -0.27 -0.05 -0.54 -0.38 -0.36 -0.15 -0.21 6.18 -0.51 6.89	57.08		5.43	4.10	5.64	1.78	ง ง	V	3.63	0.0	•		'n
1.96 0.75 0.18 -0.89 -0.23 -0.36 -0.48 -0.52 ·0.48 0.35 0.35 0.48 -0.52 ·0.48 0.35 0.35 0.48 -0.35 0.35 0.35 0.35 0.35 0.35 0.35 0.35	<u>څ</u> ،		5.04										
6.33 6.47 6.12 -6.68 -6.69 -6.16 -9.68 -9.15 -9.	98.		60.00	-6.23	-0.36	-6.48	-0.52	6.48	-6.36	-6.24	9.30	6.05	9.69 0.69
6.95 6.47 6.17 6.13 6.15 6.19 6.22 6.25 6.25 6.35 7.3 6.15 6.19 6.25 6.25 6.25 6.35 7.3 6.15 6.15 6.15 6.15 6.15 6.15 6.15 6.15	9.33		80.0-	. 6.6 3	-6.10	9.00	-0.15	-6.15	-9.24	9.6	-6.13	÷	9
5.3	96.9		6.13	6.15	6.19	9.55	9.26	9.36	6.48	6.58	9.0	S	7
	65		001										
	- 9			-0.21	10.14	-0.08	-0.27	-9.65	6.01	-6.63	-0.11	. e	9.9
6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6				. v.	61.0	8	-0.27	-8.05	9.61	9.05	-0.13	0.00	æ .
The part of the pa	9.6			₩. ₩.	6 .35	Š	0.51	8.80	80.08	96.	. S.	 	.

TABLE 2-3.

Phoneme Library for Simulated Speaker = SPBB

HERE IS THE CATALOG OF ENTRIES FOR! SPCC.SIM HEADING!

TABLE 2-4

Phoneme Library for Simulated Speaker = SPCC

3. FUNGERPPINT SIMULATION

3.1 INTRODUCTION

Fingerprints have been considered as a means to identify a person for more than 100 years. Together with voice-print, they have been known as the most reliable and difficult-to-duplicate attributes for personal identification (Table 1-1). To classify a fingerprint, minutae such as ridge endings and bifurcations (ridge branches) are usually used as basic parameters for identification. An example of such a system is the prototype system, called FINDER, developed by the Calspan Corporation for the FBI. Other parameters which are also used for classification are ridge counts, core-to-delta distances and angles, crease lengths and core-to-crease distances.

Although the aim of all automation in fingerprint classification is to produce a low-cost system that is reliable, fast and accurate, the ultimate decision regarding the selection of equipment is influenced by system environment and user needs. Basically, the application of fingerprint identification is classified into four categories: access control, transaction control, production processing, and latent print identification. This chapter introduces a method of fingerprint simulation for testing access control devices as part of our research effort in UNITRANS.

3.2 FINGERPRINT PATTERN CLASSIFICATION

One of the well-known fingerprint classification systems is the Henry system. This system consists of four major types: arches, loops, whorls, and accidentals. Each type is subdivided as follows [6]:

- 1.) Loop:
 - a.) Radial Loop
 - b.) Ulnar Loop
- 2.) Arch:
 - a.) Plain Arch
 - b.) Tented Arch
- 3.) Whorl:
 - a.) Plain Whorl
 - b.) Central Pocket Loop
 - c.) Double Loops
- 4.) Accidental

In order to have full understanding of the fingerprint patterns, the definition of some technical terms is necessary. A loop is that type of fingerprint in which one or more ridges enter on either side of the image, recurve, and terminate on the same side of the impression (Figure 3-1(a)-(b)). A radial loop is one type of loop whose ridges flow towards the thumb, and a Ulnar loop is another type whose ridges flow toward the little finger. The arch is another major type fingerprint in which the ridges enter from one side, rise in the middle, and flow out on the other side (Figure 3-1(e)-(f)). The tented arch is an arch except part of the ridges form an upward thrust at the center. A whorl is any pattern with two deltas and at least one recurring ridge which may be a spiral or any variation of a circle (Figure 3-1(c),(d),(g)). The double loop or twin loop pattern consists of two deltas and two separate loops, with separate distinct shoulders. The central pocket loop has two deltas in which one encloses the other. Accidental is the combination of all three major types mentioned before (Figure 3-1(h)).

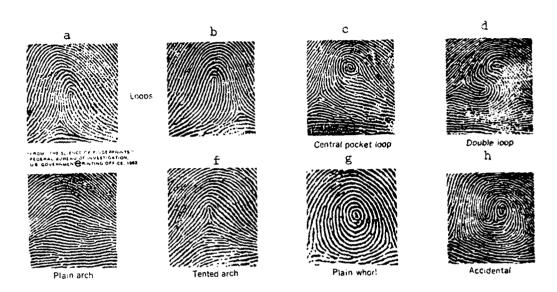


Figure 3-1 Fingerprint Patterns of the Henry System

3.3 GENERATION OF FINGERPRINT IMAGE

At the present state-of-the-art, two basic types of fingerprint identification systems are available to users: (1) those that rely on the computerized digital filtering and pattern recognition, and (2) those that use coherent optical correlation. In general, the functions of these systems can be divided into several stages: sensing or reading; image digitizing; print registration; feature extraction; and comparison and verification. For entry control purpose, sensing 'live' fingerprints instead of scanning print images from film or glass is used. Hence, two stages are under consideration in this research: (1) the computer generation of fingerprints, (2) the reproduction of 'live' fingerprints.

3.3.1 Computer Simulation

The overall approach to the computer generation of a fingerprint image involves:

- 1. Storing several 'standard' pattern types (whorl, loop, etc.) in the computer.
- 2. Setting the ridge width and interridge distances.
- 3. Deforming the pattern via topological transformations.
- 4. Super-imposing minutiae in the image.
- 5. Displaying the image for interaction with an operator who may adjust parameters (or override the computer) at each step.

Initially, three major types of fingerprints are stored in the computer as the basic patterns. They are whorl, loop, and arch; but more standard patterns can be added later, if desired.

Each basic pattern is represented by 256 x 256 pixels and is stored in 16 different files corresponding to 16 different regions of the print. The image cell (X,Y) contains logic true (.TRUE.) if it is part of the ridge, or logic

false (.FALSE.) if it is part of the sulci (valley). This is equivalent to a two level gray scale. An operator can select the pattern type, image region(s) and deform the image by a series of topological transformations (one-to-one bicontinuous mappings) to form the global fingerprint pattern. At present, we are using standard transformations of translation, rotation, stretching in the X and/or Y direction, and reflection as well as localized distortions such as produced by

localized in X direction:

$$Y' = Y + k_1 e^{-A}, (X-Xe)^2$$
(3.1)

and

localized in Y direction:

$$X' = X + k_2 e^{-A_2} (Y-Yo)^2$$
 (3.2)

In these localized deformations the parameters X_0 , Y_0 , k_1 , k_2 , A_1 , and A_2 are chosen randomly by the operator to produce the final global pattern for the fingerprint.

To impose minutia such as bifurcations on the pattern, Figure 3-2 shows a good detail. First the operator decides that a minutia is located at cell (X,Y), Figure 3-2(a). Then it is decided that the minutia is a bifurcation with a fork of length 1, and a local deformation is made to create room for the new branch, Figure 3-2(b). Finally, the new branch is introduced by changing the gray level of the appropriate cells from .FALSE. to .IRUE., Figure 3-2(c). In any event, a sophisticated and experienced operator is necessary to make a satisfactory fingerprint.

3.3.2 Physical Fingerprint Simulation

As mentioned before, access control devices require the actual physical presence of the fingerprint on the sensor glass. In order to have a complete check-up of such devices, including the testing of sensing and digitizing units, reproduction of the physical fingerprint is necessary.







Figure 3-2 Impose Minutiae on Fingerprint Pattern

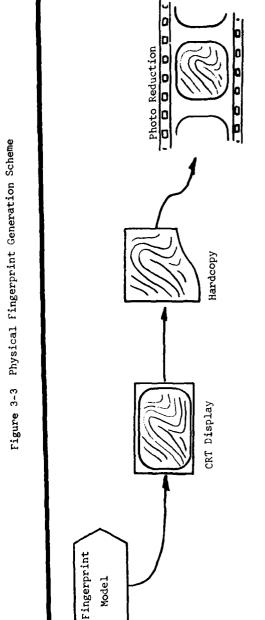
The basic procedure for physical fingerprint generation can be summarized in the following steps (see Figure 3-3).

- 1. Making a hard copy of the fingerprint image simulated by using the method described in Section 3.3.1.
- 2. Photo-reducing the image down to actual fingerprint size using a high contrast emulsion.
- 3. Photo-etching the image (photo-etch kit such as KIT 650 Photo-Etch Kit, manufactured by INJECTORALL can be used. Alternatively, the sensitized board can be made by using Photo Resist Spray manufactured by the same company).
- 4. Creating elastic compound cast (positive) from the copper clad board (negative) prepared in Step 3.

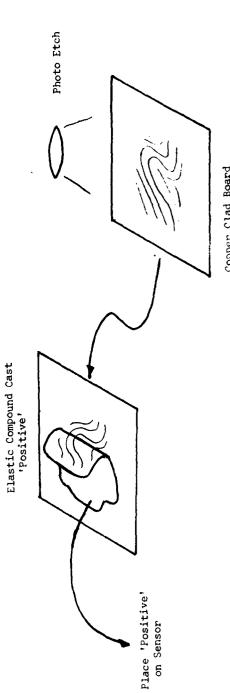
In order to generate a high quality fingerprint image by the photoetching method, timing such as exposing time and etching time are very important. But once the timing is ascertained, large production of high quality fingerprints is not a problem.

3.4 TYPICAL EXAMPLES

As mentioned before, the whole fingerprint is represented by 256 x 256 pixels and is stored in 16 different files corresponding to 16 different regions (A-P) of the print (Figure 3-4).



3-8



Copper Clad Board 'Negative'

		Center (L,J)				
	£	of Distortion	Extent of Distortion	Maximum Distortion	Distortion Angle*	Lower Bound
steps	Kegions	(Pixels)	(Pixels)	(Pixels)	(deg)	Upper Bound
Ę.	BEDGJMLO	(128, 64)	50.	7.	.0	128, 256
2	BEDGJMLO	(64, 64)	.04	10.	.0	3, 70
က	EFGHMNOP	(128, 64)	,0 t	5.	0.	3. 256
⋣	ABCDIJKL	(64, 64)	20.	.9	0.	10, 128
5	ABCDIJKL	(128, 100)	20.	-6.	.0	120, 190
9	CDGHIJMN	(64, 128)	.09	•	.06	1, 257
7	ABCDIJKL	(64, 1)	.09	-15.	.0	2, 128
8	BEDGJMLO	(64, 64)	.04	-4-	0.	3, 110

TABLE 3-1

Steps Illustrating 'Loop' Type Fingerprint Simulation for Figure 3.5

* Angle of direction of distortion measured from vertical.

L = Vertical

J = Horizontal

		Center (L,J) of Distortion	Extent of Distortion	Maximum Distortion	Distortion Angle*	Lower Bound
Steps	Regions	(pixels)	(pixels)	(pixels)	(deg.)	Upper Bound
1	BEDGJMLO	(128, 64)	30.	-5.	.0	120, 128
2	BEDGJMLO	(30, 64)	35.	-8.	0.	15, 40
3	BEDGJMLO	(128, 60)	70.	14.	0.	39, 256
tī	ABCDIJKL	(128, 2)	70.	-20.	0.	и, 190
5	ABCDIJKL	(128, 126)	70.	-20.	0.	10, 140
9	BEDGJMLO	(128, 64)	20.	-7.	0.	64, 190
7	CDGHIJMN	(64, 128)	75.	-8-	.06	116, 240
8	EFGHMNOP	(128, 64)	30.	-6.	45.	128, 370
6	ABCDIJKL	(64, 64)	20.	-9-	-45.	-100, 120
10	BEDGJMLO	(128, 64)	20.	-5.	0.	123,132

TABLE 3-2

Steps Illustrating 'Arch' Type Fingerprint Simulation for Figure 3.6

*Angle direction of distortion measured from vertical.

L = Vertical J = Horizontal

1 BEDGJMLO (128, 64) 50. 13. (2 BEDGJMLO (128, 64) 506. (4 ABCDIJKL (128, 64) 30. 7. (5 EFGHMNOP (128, 110) 7410. (6 EFGHMNOP (64, 64) 50. 10. (7 CDGHIJMN (64, 128) 606. 9(8 BEDGJMLO (128, 64) 20. =8. (9 ABCDIJKL (128, 64) 407. (Steps	Regions	Center (L,J) of Distortion (pixels)	Extent of Distortion (pixels)	Maximum Distortion (pixels)	Distortion Angle [*] (deg.)	Lower Bound & Upper Bound
BEDGJMLO (128, 64) 50. -6. ABCDIJKL (128, 64) 30. 7. ABCDIJKL (128, 60) 40. -5. EFGHMNOP (64, 64) 50. 10. EFGHMNOP (64, 64) 50. 10. CDGHIJMN (64, 128) 60. -6. BEDGJMLO (128, 64) 20. =8. ABCDIJKL (128, 1) 40. -7.	1	BEDGJMLO	(128, 64)	50.	13.	0.	125, 256
ABCDIJKL (128, 64) 30. 7. ABCDIJKL (128, 60) 40. -5. EFGHMNOP (128, 110) 74. -10. EFGHMNOP (64, 64) 50. 10. CDGHIJMN (64, 128) 60. -6. BEDGJMLO (128, 64) 20. =8. ABCDIJKL (128, 1) 40. -7.	5	BEDGJMLO	(128, 64)	50.	-9-	0.	3, 110
ABCDIJKL (128, 60) 40551010101010101010	е	ABCDIJKL	(128, 64)	30.	7.	0.	10, 140
EFGHMNOP (128, 110) 74. -10. EFGHMNOP (64, 64) 50. 10. CDGHIJMN (64, 128) 60. -6. BEDGJMLO (128, 64) 20. =8. ABCDIJKL (128, 1) 40. -7.	±	ABCDIJKL	(128, 60)	*0†	-5.	0,	31, 256
EFGHMNOP (64,64) 50. 10. CDGHIJMN (64,128) 60. -6. BEDGJMLO (128,64) 20. =8. ABCDIJKL (128,1) 40. -7.	ī	EFGHMNOP	(128, 110)	.47	-10.	0,	3, 256
CDGHIJMN (64, 128) 606. BEDGJMLO (128, 64) 20. =8. ABCDIJKL (128, 1) 407.	ဝ	EFGHMNOP	(64, 64)	.05	10.	0.	10, 128
BEDGJMLO (128, 64) 20. =8. ABCDIJKL (128, 1) 407.	7	CDGHIJMN	(64, 128)	•09	-9-	•06	1, 230
ABCDIJKL (128, 1) 407.	ω	BEDGJMLO	(128, 64)	20.	-8.	0.	130, 140
	6	ABCDIJKL	(128, 1)	.04	-7.	0.	2, 256

TABLE 3-3

Steps Illustrating 'Whorl' Type Fingerprint Simulation for Figure 3-7

* Angle of direction of distortion measured from vertical.

L = Vertical
J = Horizontal

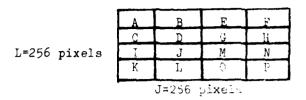


Figure 3-4. Regions Assignment for Fingerprint Simulation

By entering a command string consisting of pattern type, region for distortion, center of distortion, lower bound and upper bound, etc., a desired type of fingerprint can be generated.

For the purpose of illustration, the procedure involved in such simulation is summarized in Table 3-1 to 3-3 and Figure 3-5 to 3-7. From the results shown, it is interesting to note that after the first few steps each pattern generated can actually be treated as a new simulated print. That is to say, a large data set can be assembled once the first complete and satisfactory print is formed. The resulting copper clad board with fingerprint image from photo-etching process is also shown in Figure 3-8.

3.5 FURTHER DEVELOPMENT

So far the method we described is only semi-automatic. The locations and occurrence of minutia are totally up to the operator. However, the print can be simulated statistically by using the Monte Carlo technique we discussed before. Since the percentage of fingerprints falling into the loop, whorl, or arch is well known the statistics of the random draw of fingerprint types can be set. The minutia can also be added by assuming that they are uniformly distributed within the print. Suppose the expected number of minutia in a given print is 70, then for each cell (X,Y) the computer will decide with a probability $70/n^2$, where n is the grid size, and (X,Y) is the location of a minutia. If (X,Y) = .TRUE. (i.e., the location of the minutiae is on a ridge) a random (probability 1/2) decision will be made as to whether it is a ridge ending or a bifurcation. If (X,Y) = .FALSE., a random decision will be made as to whether it is the ridge ending of a short ridge or a rudimentary ridge.



BASE PATTERN



Step 1



Oten I



Oten 3



1.30 5



Step



Step 6

Ti to 3-5 Loop Type Fingerprint Simulation

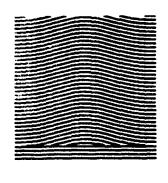




Step 7

Step 8

Figure 3-5 Loop Type Fingerprint Simulation (Cont.)



BASE PATTERN

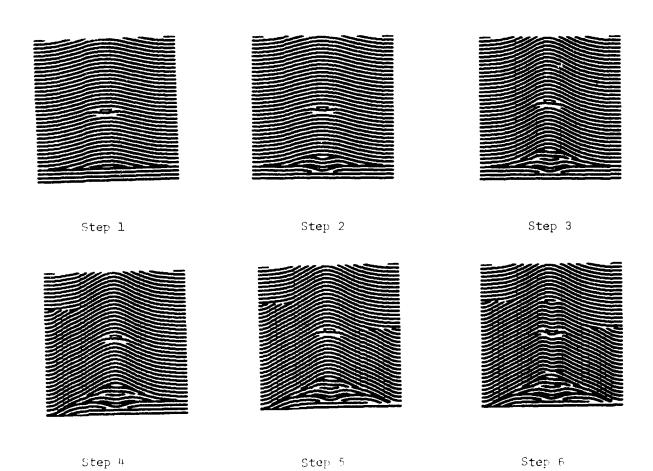
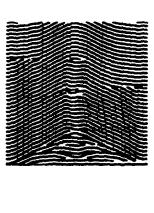
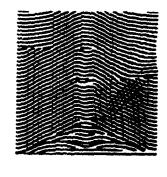


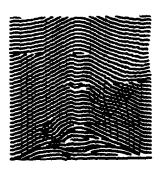
Figure 3-6 Arch Type Fingerprint Simulation



Step 7



Step 8



Step 9



Step 10

Figure 3-6 Arch Type Fingerprint Simulation (Cont.)



BASE PATTERN

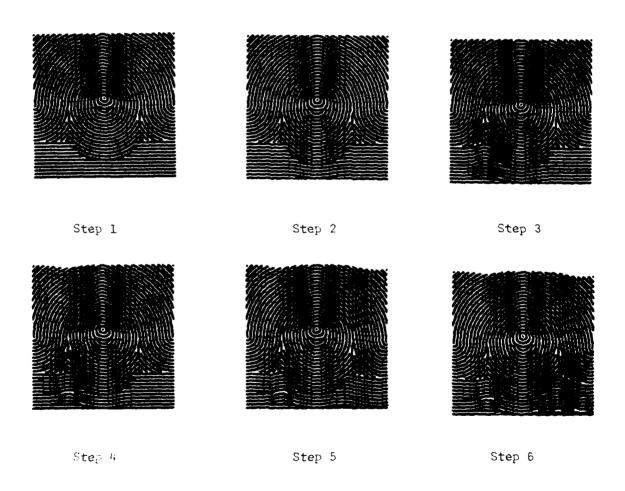
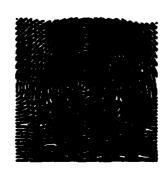
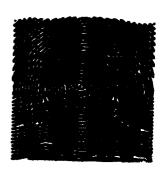


Figure 3-7 Whorl Type Fingerprint Simulation







Step 7

Step 8

Step 9

Figure 3-7 Whorl Type Fingerprint Simulation (Cont.)

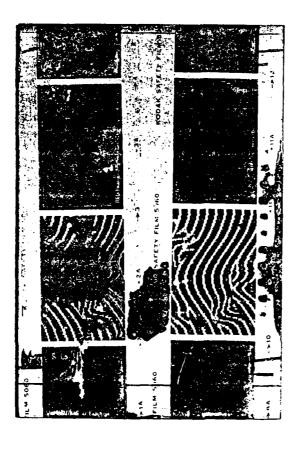


Figure 3-8. Copper Clad Board With Fingerprint Images Taken in Different Exposures. (Images have not been photo-reduced.)

Similarly, random decisions can be made to determine length of a break in a ridge, length of a ridge created by a bifurcation and whether it rejoins the main branch (creating an enclosure), and the lengths of short ridges and rudimentary ridges.

4. SIGNATURE SIMULATION

4.1 INTRODUCTION

In signing one's name the tip of the pen describes an arc through space. From the beginning of the signature when the pen first touches the paper to the final stroke, a very mechanical sequence of motions are followed. When the pen is down on the paper, three functions of time are necessary to describe its history, x(t) and y(t), the path across the paper in Cartesian coordinates, and p(t), the pressure on the paper. When the pen is lifted, as occurs in the break between names and in some letters such as 'i', the motion of the pen tip occurs in three dimensional space so that pressure must be replaced with z(t), the height above the paper. Thus p(t) and z(t) are complementary. Only one is needed at any given time: p is nonzero only when z(t) is zero and conversely. An autograph is the residual trace of this three-dimensional motion as it contacts the two-dimensional space of the writing surface. Effects of pressure and duration of each segment of the curve become hidden or disappear altogether.

An autograph, by its existence in the plane of a writing surface, partakes of the quality of an image. The signature, by which we mean the act of signing, partakes of the qualities of a trajectory in space. That is, time and velocity are important variables. The purpose of this section is to consider some potential models for handwritten signatures. We shall see that the two different models for signature are based on the different properties of images and trajectories.

4.2 MODELS FOR SIGNATURES

Consider the autographs of Figure 4-1. In attempting to find a model for signatures, we need first to determine what is repeatable. These five autographs appear to casual inspection to be the product of a single author, as indeed they are. What is it about them that is constant or repeatable? When these autographs are examined in detail, they are different in each detail, the center loop of the capital 'R', the exact position of the period, the shape of the tail on the final 'n', and so forth. Just as two leaves from a tree are recognized as very similar in total structure, yet no two are identical, these autographs in Figure 4-1 are recognized as being of the same hand, even though each is unique. Evidently the human brain does an excellent job of subconsciously identifying the similarities.

The psychologist, Werner Wolff, has discussed the symmetries which are often found in autographs [7]. In so doing, he is emphasizing the qualities of a signature which are image-like. He describes, for example, how experiments show that in autographs containing a first and last name the lengths are frequently proportional with small integer constants of proportionality. He often observes balanced symmetry about the period of a central initial. Or, another type of regularity concerns the "measuring stroke" in which one prominent stroke of the signature serves as a scale by which the sizes of other parts may be predicted. In all these cases the size of the autograph may vary, but the symmetries and proportionalities are preserved.

The autograph may almost be considered a picture which the signer is drawing. The pictorial qualities are emphasized. One would be led to suggest that the appearance of the signature, the autograph, is what determines the motion of the signer's hand. That is, the signer visualizes what the signature is to look like and moves his hand to reproduce the visualized image. Some signatures even contain pictograms or leitmotifs such as crosses, musical symbols, palettes, flags, swords, guns, flowers, and birds [7]. In

Mark R. Mahan Mark R. Mahan Mark R. Mahan Mark R. Mahan

Figure 4-1 Five Repetitions of a Signature by the Same Individual

the case of Ludwig II of Bavaria, a swan, symbol of his house can be seen in his autograph. (See Figure 4-2.)

A rather different perspective is gained by considering the motion of the signature. In this model the signature is a sequence of muscle motions which have become automatic. The signing of a name is like executing the steps of a complicated dance, and any pictorial aspects are incidental. The trajectory described by the pen point is the major factor.

We have described two generically different conceptions of a signature, the pictorial, visual, image-like, static versus the choreographic, kinetic, dynamic. The distinction is rather important in determining the type of deviations one expects to find in repeated signatures. For example, returning to Figure 4-1, we observe that the three capital letters all have the same initiating stroke. Suppose an error in hand motion causes the stroke on the 'M' to be enlarged. Then if the signer is attempting to maintain pictorial symmetry, he might attempt to repeat the erroneous stroke on the 'R' and 'N'. If, on the other hand, the trajectory model is a better description, we find no such attempt to symmetrize or balance. Disturbances away from the ideal would be much more localized at the level of individual pen strokes.

An interesting and simple experiment which can suggest the relative importance of the trajectory and pictorial aspects of a signature is to compare normal signatures executed with eyes open and those signed with eyes closed. In this context, compare Figures 4-3a and 4-3b, 4-4a and 4-4b, and 4-5a and 4-5b. Surprisingly little deterioration is caused by closing the eyes. This fact suggests that the trajectory model of a signature is more accurate. Obviously the pictorial quality of the signature is important, but Figures 4-3 and 4-5 suggest that the sequence of motions is very mechanical. That is, the appearance of the signature, the autograph, has determined through a demand for symmetry, symbolism, and proportionality what sequence of motions will be learned in a very general way. The actual storage of the signature in our memories, however, would seem to have very few pictorial

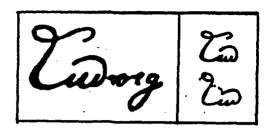


Figure 4-2. Signature of Ludwig II of Bavaria, Which Contains a Swan, Symbol of His House.

Non Sheh Non Sheh Non Sheh Non Sheh Non Sheh

Nou Show Non Show Non Show Non Show Non Show

Figure 4-3a Five Signature Repetitions - Eyes Open

Figure 4-3b Five Signature Repetitions - Eyes Closed

Rebord J. Jackson

Rebord J. Jackson

Rebord J. Jackson

Rebord J. Jackson

Figure 4-4a Five Signature Repetitions - Eyes Open

Rubord D. Jackson

Figure 4-4b Five Signature Repetitions - Eyes Closed

James 195 James 195 James 199

Figure 4-5a Five Signature Repetitions - Eyes Open

Figure 4-5b Five Signature Repetitions Eyes Closed

James Cogge

James & Cong-

James 1895

qualities, but rather to be a sequence of muscle movements. We shall therefore proceed to build up a model for signatures based on the trajectory of the pen point.

Note that there is an element of the signature which the trajectory representation of x, y, and p cannot describe, namely, the width of the stroke, particularly as it varies along the signature. This is an especially prominent feature when a signature is executed with an instrument having asymmetrical tip such as a fountain pen: The flat tip of a pen can make broad or narrow strokes depending upon the angle at which the nib addresses the paper. However, with the advent of symmetrical pens such as the ball point and felt tipped, this dimension of a signature has virtually disappeared.

We shall assume in the following that a signature is described by x(t), y(t), and p(t), (or z(t)). These signals may be represented by their sampled time series, \times_{α} , y_{α} , p_{α} , where the samples are at regular closely spaced intervals. Let us now concentrate on the two spatial coordinates x and y in the plane of the paper. When p is non-zero, the pen is down and all motion is within the x-y plane. To emphasize the trajectory aspect of the signature, the signature can be encoded as \dot{x}_{α} and \dot{y}_{α} , where \dot{x}_{α} is the x component of the velocity near time α and is given by $x_{\alpha+1} - x_{\alpha}$. Similarly for y. Thus, the signature is encoded as a sequence of small displacement vectors (see Figure 4-6).

We would like to illustrate the above encoding scheme with the signature shown in Figure 4-1, which was sampled at 50 Hz on a Tektronix 4953 Graphics Tablet with a spatial resolution of 100 l.es/inch. The components of velocity are given in Table 4-1 for the capital letter 'M'. The values of $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are plotted against time in Figure 4-7. We note that there is a good deal of correlation between the two curves. The reason for this is obvious when the vectors of Table 4-1 are plotted in polar form as in Figure 4-8. The speed of the trajectory, $\mathbf{v} = (\mathbf{x}^2 + \mathbf{y}^2)^{1/2}$ is a relatively featureless function with its details being not much larger than the noise. The direction of

Jak ark

Figure 4-6. Signatures may be encoded as a sequence of velocity vectors.

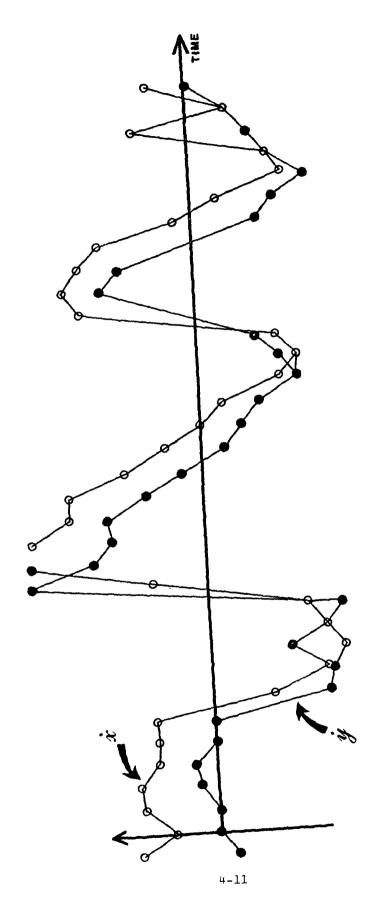


FIG. 4-7. LETTER 'M' \hat{x} and \hat{y} curves.

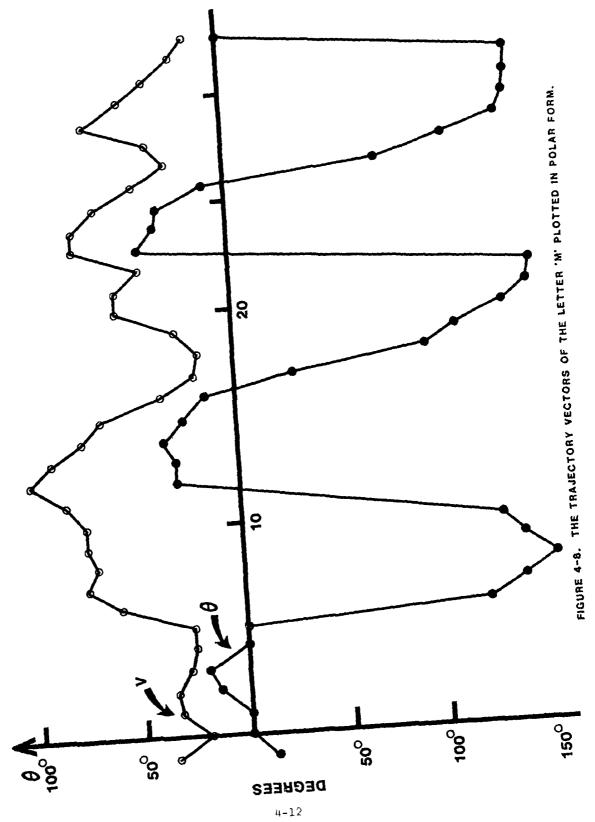


Table 4-1

Velocity vectors from the trajectory which form the capital 'M' of Figure 1. The units per .02 seconds, or .5 inches.

x	j	v	θ
35	-8	36	-13
18	0	18	0
34	0	34	0
35	8	36	13
26	9	28	19
26	0	26	0
27	0	27	0
-27	-52	59	-117
-52	- 53	74	-134
-61	-3 5	70	150
-52	-52	74	-135
-44	- 61	75	-125
26	79	83	72
87	52	101	31
79	43	90	29
61	44	7 5	36
61	26	66	23
35	9	36	ļ4
17	-9	19	-28
0	-17	17	-90
-8	-26	27	-107

TABLE 4-1. (Continued)

×	у	v	0
-35	-44	56	~129
-44	-35	56	-141
-35	-26	44	-143
53	52	74	44
61	44	75	3€
52	35	63	34
43	9	44	11
9	-27	28	-72
-9	-35	36	-104
-43	-52	€7	-130
-35	-35	49	-135
-26	-26	38	-135
-17	-17	24	-1 35 ·
17	0	17	0

motion, Θ , however, shows the strong discontinuities of almost 180° which result from the cusps at the bottom of the 'M'. To first order, the speed of trajectory is constant. Only the direction of motion changes. This accounts for the correlation in \mathbf{x} and \mathbf{y} since there is really only a single independent function, Θ . It would therefore seem that the polar representation is superior.

A little thought shows that the existence of cusps, sharp reversals in angle, are a common feature of script. Figure 4-9 gives the letters of the Palmer system of handwriting and indicates the number of cusps in each. Since these cusps are easily detected in the 0 curves and since the strokes between cusps are relatively uncomplicated, a good representation of a signature would begin by dividing the curve 0 into segments having boundaries at the cusps. The curve segments would then be represented by a small set of basis functions. The velocity and pressure curves on these intervals would also be expanded in orthonormal functions.

The signature would then be modelled as N segments, each of duration T_i , for $i=1,2,\ldots,N$. The segments occur at points in time when the pen is raised or lowered (a transition from pressure recording to elevation recording) or at cusps. The θ , v, and p time segments at interval i are expanded with a set of basis functions,

A signature is given by the set

$$T_1, a_1^{(1)}, a_2^{(1)}, \dots, a_q^{(1)}, b_1^{(1)}, b_2^{(1)}, \dots, b_r^{(1)}, c_1^{(1)}, \dots, c_p^{(1)}$$
 $T_2, a_1^{(2)}, \dots, c_p^{(N)}$
 $T_N, a_1^{(N)}, \dots, c_p^{(N)}$

α 1/2	1/1	<i>C</i> 0/1	d 1/2	0/0	1/1	1/2
L 1/1	<i>i</i> 1/1	Ť ₁	£ 2/2	<u>L</u> 0/0	m 2/2	M 1/1
<i>O</i> 1/2	2/2	g 1/2	N 1/1	2/2	<i>t</i>	20 2
٧ 1/1	M 3/3	0/0	<i>y</i>	36 1/1		

Figure 4-9 Letters written according to Palmer system. Below the numerator of the fraction gives number of cusps found in the letter in unconnected writing, the denominator in for connected writing.

It will be important to recognize the possibility of cross-correlation between the a_j , b_j , and c_j coefficients.

Alternatively, it may prove useful to model the behavior of the segments between cusps in terms of linear predictive coefficients, as described in Section 2.

Having rejected the pictorial aspect of a signature, we now compromise our viewpoint by acknowledging the likelihood that certain distortions characteristics of images may occur in a signature. In particular, a signature need not be rotationally aligned to permit its recognition. A uniform rotation of the signature can be accommodated as a constant offset on the θ curve. Changes of signature scale can likewise be expected. It seems probable that the scaling can be described as a stretching of the trajectory variable in time. That is $\theta(t)$ would be replaced by $\theta(\lambda t)$, where λ is a scale factor.

In any case, a literature search has revealed very little information on modelling of signatures at the level of detail which is required by UNITRANS. It would appear useful to collect some data with the Tektronix Graphics Tablet to verify the model proposed in this section. Also, should the use of basis functions to represent strokes between cusps be satisfactory, the experimental data could serve to suggest some possible choices for the functions.

5. COMPARISON OF UNITRANS WITH VADABS SYSTEM

5.1 INTRODUCTION

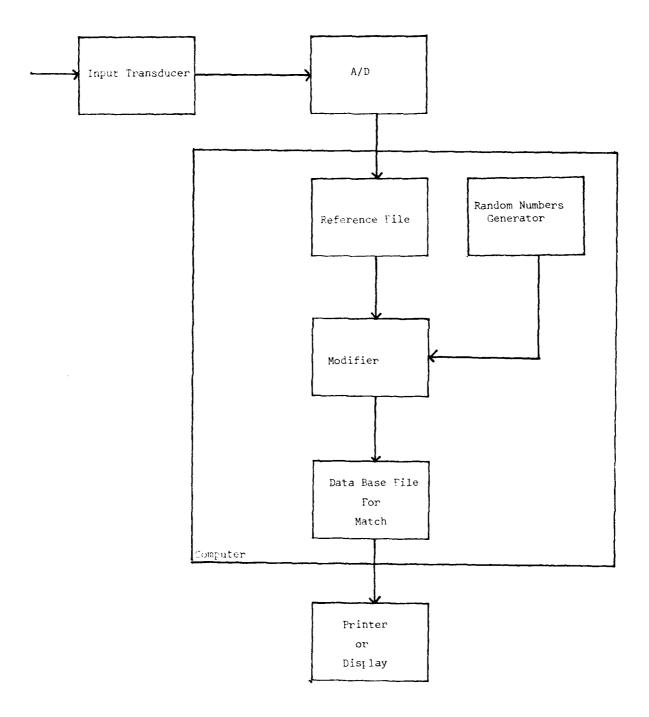
As mentioned in Section 1 collecting sufficient data for device testing is a difficult problem and it is not surprising to find that another technique, VADABS (Variable Data Base), has been proposed [9] to perform the same function as UNITRANS. Therefore, it is worthwhile to discuss the VADABS concept and our different approach to the UNITRANS. We will first discuss the VADABS system and its limitations, followed by the presentation of our UNITRANS system and its differences compared with the VADABS.

5.2 VADABS SYSTEM AND ITS LIMITATION

The VADABS system attempts to collect a relatively small sample of actual data in digital form and from this construct a reference file of real data. The input signals of the reference file would then be modified to form a larger data base as shown in Figure 9-1.

Several limitations to the VADABS system are then apparent and are summarized as follows:

- 1. It cannot be used to test an entry control system which is all or partially analog. For example, an optical fingerprint recognition system or an entry control system with a person in the loop.
- 2. Even for a digital system only a portion of the entry control system can be tested. The input transducer, A/D converter, or any feature extraction device cannot be tested.



Playing 5-1 VADABS System Concept

3. It cannot use external information about the rest of the input signal population outside that represented by the reference file.

In addition, there are some points that need clarification:

- 1. Do the reference and expanded data base files contain raw data or input signal features?
- 2. How does the modifier work?
- 3. Can the same modifier perform the same modifications on input signals regardless of whether they are speech, fingerprints, signatures, or whatever?

A partial but unsatisfactory answer for question two is suggested, namely that random noise, in accordance with known information on the standard deviation, be added to the input signal. Also it is suggested that the input signals in the reference file be intermixed. In the first instance, the standard deviation on the input signal cannot be known unless enough samples have been collected to estimate it, in which case nothing new is created in the data base file. To produce something new it would be necessary to put large random fluctuations on the signal and it is hardly conceivable that the result could represent any real input signal. Small random noise could be considered as variations on an individual's input, but to know the standard deviations involved implies that there are already enough samples for that individual.

In the second case, it is not clear how the input signal can be intermixed to produce anything resembling a real input.

Finally, and perhaps more damaging than the previous criticisms, there is a logical objection. It is proposed that the modifer changes the input signal to obtain a new signal that is "similar but different enough to be utilized as test data". This implies that:

- 1. There is some objective measure of when two signals are "different enough", or
- 2. An operator subjectively assesses when two signals are "different enough", or
- 3. system being tested.

In case one, one already has a perfect entry control system against which one is measuring the effectiveness of others which aren't really needed. In case two, the entry control system being tested can be made to look as good or as bad as the operator chooses. In case three, the evaluation has little or no meaning since "different enough" can only mean "different enough" for the system to work.

5.3 UNITRANS APPROACH

PAR, recognizing the problems associated with simulating a data base which would be useful for testing entry control systems, has taken a somewhat different approach. The considerations that the data base be useful for testing analog as well as digital systems and that it test all parts of the system dictated that it be possible to generate inputs that could be fed directly into the input transducer (i.e., sound waves, artificial fingerprints, etc.).

The difficulties inherent in trying to modify existing samples to obtain realistic data are circumvented by extracting features from the signal which are adequate to synthesize the necessary inputs and are such that the overall distribution of these features as well as distributions for individuals can be estimated based on our own experimentation and data collected by other researchers. From these distributions, then, the computer can sample and produce a realistic data base of features from which signals to the input transducer of the device under test can be synthesized.

6. RECOMMENDATIONS

The success achieved in simulating the attributes fingerprint and signature was in direct proportion to the percentage of the effort expended on them: speech reproduction is the most well developed of the techniques, fingerprint simulation seems moderately successful, while signature simulation was carried only to the level of concept formulation.

Simulated Speech is reasonably natural sounding. Moreover, a complete system for reproducing simulated speech could be assembled at moderate cost, especially if simulated speech were only one mode of operation of a general system capable of reproducing both a data base of actual speakers and the simulated speakers. This is, of course, the philosophy adopted in this project. As described in Section 2, the simulated speech is produced by the same system which is employed for constructing new utterances by actual data base speakers.

A system for reproducing simulated and actual speakers would consist of a small computer with disk and console, an analog-to-digital converter, a digital-to-analog converter, and audio subsystem of microphone and loudspeaker. Such a speaker synthesis unit would be simple and easy to use.

Speaker synthesis might be used in testing of speaker verification devices, as originally envisioned in this project. The fixed data base of speakers would provide a base line against which any new device would be tested. The simulated speakers, on the other hand, could be used to probe the limits of the verification device. The device performance could be measured for various deviations away from the nominal population mean. Recall that the simulated cross-sectional areas, A, are generated according to

 $A = \overline{A} + x\sigma$.

where \overline{A} is the population mean, x is a Gaussian random number, and σ is the population standard deviation. Let k be the standard deviation of the distribution from which x is drawn. Then, the device performance could be tested as a function of k to show its response to extremes in the population.

It is also possible to use a speaker synthesis system to est and exercise a speech recognition device. In this case one would be interested in showing the device performance on a diverse group of words and speakers. For speech recognition systems which require training on the speaker's voice, it would be possible to train the device and then, using UNITRANS, add a measurable distortion to observe the tolerance. A speaker synthesis unit might also be useful in designing the word lists to be employed in an operational speech recognition system. One could ascertain in advance whether the word list attained good performance over a spectrum of speakers.

In summary, it is possible to envision a self-contained, reasonably simple speaker synthesis unit which would serve as a laboratory tool in a variety of capacities. We recommend that RADC procure a speaker synthesis unit and evaluate its utility in a speech processing laboratory environment.

By comparison, fingerprint simulation is more of a technique than a device. The production of simulated rubber fingerprints requires many steps. It would appear much simpler and cost effective to collect fingerprint samples from a set of real people. Furthermore, a simulated fingerprint will be of less general utility than simulated speech. Fingerprints can only be used for testing identity verification devices, whereas speech is used both for identity verification and communication. Thus a speaker synthesis unit would serve to test two classes of devices, those that recognize speakers and those that recognize speech, while fingerprint synthesis would test only one class.

Fingerprint simulation would be much more useful if requirements of the application were such that the digital fingerprint image was never output. For example, suppose one were simulating an access control system with a

fingerprint verification device, and the simulation was entirely in software. Then a subroutine to generate simulated fingerprint patterns would be useful and the UNITRANS technology would serve well. Only in this context do we see fingerprint simulation being a cost effective technique.

As we have stated before, signature simulation has been developed only to the level of a concept. In conclusion, we can state that speaker synthesis may be performed by a <u>device</u> which we believe would be of immediate utility to RADC, fingerprint synthesis is a <u>technology</u> which may have limited utility depending upon RADC future projects, and signature simulation is a <u>concept</u> only, awaiting future requirements.

REFERENCES

- [1] Swonger, C.W., "Applications of Fingerprint Identification Technology to Criminal Identification and Security System," presented at 1st International Conference on Electronic Crime Countermeasures, Edinburgh, Scotland, July 18-20, 1973.
- [2] Forsen, G., Nelson, M., and Staron, R., "Personal Attributes Authentication Techniques," Final Technical Report, Contract Number F30602-76-C-0368, Report Number RADC-TR-77-333, Rome Air Development Center, Griffiss AFB, NY, A047645.
- [3] LeBoss, B., "Speech I/O Is Making Itself Heard," Electronics, pp. 95-105, May 22, 1980.
- [4] Iversen, W.R., "Military Has Its Eye on Speech," Electronics, pp. 93-94, June 5, 1980.
- [5] Markel, J.D. and Gray, A.H., Jr., "Linear Prediction of Speech," Springer-Verlag (1976).
- [6] "The Science of Fingerprints," Federal Bureau of Investigation, U.S. Government Printing Office, 1963.
- [7] Wolff, Werner, "Diagrams of the Unconscious," Grune and Stratton, New York, 1948.

- [8] Shohara, N., "Data Collection Analysis and Test," Final Technical Report, Contract No. F30602-79-C-0176, Rome Air Development Center, Griffiss AFB, NY. (To be published.)
- [9] Personal Communications with Jim Maier from Rome Air Development Center, Griffiss AFB, NY.

MISSION of Rome Air Development Center

RADC plans and executes research, development, test and selected acquisition programs in support of Command, Control Communications and Intelligence (C^3I) activities. Technical and engineering support within areas of technical competence is provided to ESP Program Offices (POs) and other ESD elements. The principal technical mission areas are communications, electromagnetic guidance and control, surveillance of ground and aerospace objects, intelligence data collection and handling, information system technology, ionospheric propagation, solid state sciences, microwave physics and electronic reliability, maintainability and compatibility.

MANGRANGHANGHANGHANGHANGHANGHANGH

